Network-based Approach to Identify Prognosis-related Genes in Tamoxifen-treated Patients With Estrogen Receptor-positive Breast Cancer

Yanyan Wang
Zhejiang University School of Medicine
Xiaonan Gong
Zhejiang University School of Medicine
Yujie Zhang (zycx@zju.edu.cn)
Zhejiang University School of Medicine Second Affiliated Hospital https://orcid.org/0000-0001-8164-6425

Research

Keywords: estrogen receptor-positive breast cancer, tamoxifen,weighted gene co-expression network analysis, biomarker

Posted Date: July 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-39426/v1

Version of Record: A version of this preprint was published at Bioscience Reports on September 1st, 2021.
See the published version at https://doi.org/10.1042/BSR20203020.

Abstract

Background: The estrogen receptor (ER) antagonist tamoxifen is the most commonly used endocrine therapy for ER-positive breast cancer. However, tamoxifen resistance remains a major cause of cancer recurrence and progression. Here, we aimed to identify hub genes involved in the progression and prognosis of ER-positive breast cancer following tamoxifen treatment.

Results: Microarray data (GSE9838) for 155 tamoxifen-treated primary ER-positive breast cancer samples were obtained from the Gene Expression Omnibus database. In total, 1706 differentially expressed genes (DEGs) associated with tamoxifen resistance, including 859 upregulated genes and 847 downregulated genes, were identified. These DEGs were mainly enriched in functions such as protein targeting to the ER and pathways such as ribosome and oxidative phosphorylation.Weighted correlation network analysis (WGCNA) clustered genes into 13 modules, among which the tan and blue modules were the most significantly related to prognosis. From these two modules, we further identified three prognosis-related hub genes (*GRSF1*, *MAPT*, and *REC8*) via survival analysis. High expression of*GRSF1* predicted poor prognosis, whereas *MAPT* and*REC8*indicated favorable survival outcomes in all patients with breast cancer and in patients with ER-positive breast cancer based on The Cancer Genome Atlas database. These hub genes were further verified by reverse transcription quantitative polymerase chain reaction.

Conclusion: Our findings established novel prognostic biomarkers to predict tamoxifen sensitivity, which may facilitate individualized management of breast cancer.

Background

Breast cancer remains the most prevalent cancer in women, accounting for 14% of cancer-related death worldwide[1]. Multidisciplinary synthetic therapies, including surgery, chemotherapy, hormone therapy, radiotherapy, and targeted therapy, have greatly improved the efficacy of treatment in patients with breast cancer[2]. However, nearly one-third of patients with primary breast cancer progress to advanced-stage cancer after resection of the primary tumor, and once breast cancer progresses to distant metastasis, the disease is considered incurable[3].

Based on gene expression profiling techniques, breast cancer can be classified into different molecular subtypes (luminal A, luminal B, basal, human epidermal growth factor receptor 2 (HER2) overexpressing, and normalcancers), which display distinct etiologies, responses to treatment, and prognoses[4]. Estrogen receptor (ER)-positive (luminal A and luminal B) breast cancers are the most common subtypes and can be treated using hormone-based therapy. Tamoxifen is a first-generation selective ER modulator that competes with estradiol for binding to the ER, thereby antagonizing the effects of estrogen and inhibiting the growth and proliferation of tumor cells[5]. The administration of tamoxifen greatly minimizes the risk of recurrence of ER-positive breast cancer, particularly for premenopausal women[6]. Unfortunately, approximately 40% of ER-positive patients are not sensitive to tamoxifen treatment and exhibit innate or acquired resistance phenotypes [7,8]. Multiple mechanisms contribute to the development of tamoxifen

resistance, including mutation or epigenetic modification of the *ER* gene and activation of alternative growth pathways [9, 10]. To date, the exact mechanisms of tamoxifen-resistance in breast cancer are still largely unknown, and few effective and reliable prognostic biomarkers have been developed to monitor tamoxifen efficiency. Thus, more studies are still needed to elucidate the molecular mechanisms of tamoxifen resistance, particularly to identify novel potential genes for monitoring treatment efficacy and predicting prognosis.

Currently, co-expression analysis has emerged as a powerful technique for mining gene expression profiles in various cancers. As an effective bioinformatics approach, weighted gene co-expression network analysis (WGCNA) was originally developed to explore the correlations among genes and to identify candidate biomarkers associated with clinical parameters[11-13]. In breast cancer, several studies have utilized WGCNA to identify hub genes closely related to clinicopathological traits (e.g., tumor size, grade, and molecular subtypes) and survival outcomes. For example, Tang *et al.* found that elevated expression of ASPM, TTK, and CDC20 conferred a poorer prognosis in breast cancer[14], and Jiang *et al*. identified six hub genes (*CA12*, *FOXA1*, *MLPH*, *XBP1*, *GATA3*, and *MAGED2*) that could serve as biomarkers for the prediction of better chemotherapeutic responses and favorable survival in patients with breast cancer[15].

Accordingly, in this study, we performed an integrated analysis based on WGCNA to identify novel biomarkers associated with tamoxifen resistance in patients with breast cancer. Additionally, the expression levels of candidate hub genes were examined by in vitro experiments in tamoxifen-resistant and -sensitive breast cancer cell lines to validate the robustness of the results.

Results

Identification of differentially expressed genes and functional annotation

According to the thresholds of $|\log_2 FC| > 1$ and $P < 0.01$, we identified 1706 differentially expressed genes (DEGs), including 859 upregulated genes and 847 downregulated genes, between tamoxifen-sensitive and -resistant breast cancer samples. A volcano plot of the DEGs is shown in Figure 1.

In Gene ontology (GO) enrichment analysis, upregulated genes were significantly enriched in various biological processes, such as establishment of protein localization to the endoplasmic reticulum and protein targeting to the ER. The downregulated genes were mainly enriched in signal release and positive regulation of hormone secretion (Table 1). In Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of upregulated DEGs, ribosome and oxidative phosphorylation were the most significantly enriched pathways. The downregulated genes were mainly enriched in neuroactive ligand-receptor interaction and the Janus kinase/signal transducer and activator of transcription (JAK-STAT) signaling pathway (Table 2).

Weighted co-expression network construction and key module identification

In our study, 28 abnormal samples were excluded (Figure 2A). The value of β = 6 was selected as the soft-thresholding power to achieve a relatively scale-free network, which was closer to the real biological network state (Figure 2B-C). We then identified 14 modules via average linkage hierarchical clustering. The DEGs in gray were not included in any module; therefore, we did not perform any functional analysis of the DEGs in gray (Figure 2D). Of these modules, the tan module showed obvious positive correlations with relapse, distant metastasis, and death. Additionally, a significant negative correlation was found between the blue module and poor prognosis (Figure 3). Hence, tan and blue modules may play essential roles in the biological processes of breast cancer tamoxifen-resistance. Thus, these modules were chosen as modules most related to disease progression for further analysis.

Identification of hub genes in the tan and blue modules

Hub genes have high connectivity within clinic-related modules and tend to play critical roles in the molecular mechanisms of tamoxifen resistance. Therefore, we next used Cytoscape to visualize hub gene networks in the tan and blue modules. As shown in Figure 4, 38 and 50 genes with the highest intramodular connectivity in the tan and blue modules, respectively, were screened out.

Identification of prognosis-related hub genes

To further explore the effects of these hub genes on survival in patients with breast cancer, we conducted survival analysis of 88 hub genes based on The Cancer Genome Atlas (TCGA) data. High expression of three hub genes, i.e., G-rich RNA sequence binding factor 1 (*GRSF1*), cytochrome c oxidase subunit 7B (*COX7B*), and chaperonin containing TCP1 subunit 8 (*CCT8*), in the tan module were all significantly associated with poor survival outcomes, whereas microtubule associated protein tau (*MAPT*) and REC8 meiotic recombination protein (*REC8*) in the blue module all predicted better prognosis in breast cancer when overexpressed (Figure 5).

Validation of hub genes in patients with ER-positive breast cancer

In addition to survival analysis based on all patients with breast cancer, we further investigated whether significant associations still existed between the five hub genes (*GRSF1*, *COX7B*, *CCT8*, *MAPT*, and *REC8*) and ER-positive breast cancer. The results revealed that three of the five genes, i.e.,*GRSF1*, *MAPT*, and *REC8*remained significantly associated with survival outcomes in patients with ER-positive breast cancer ($P < 0.05$; Figure 6A-C).

Validation of hub genes in vitro

Reverse transcription quantitative polymerase chain reaction (RT-qPCR) was then performed to further validate the expression of candidate hub genes in tamoxifen-sensitive MCF-7 and tamoxifen-resistant MCF-7/TAM cell lines. The RT-qPCR results suggested that the expression level of *GRSF1* was higher in tamoxifen-resistant cells, whereas *REC8* and *MAPT* expression was higher in MCF-7 cells than in parental MCF-7/TAM cells (Figure 6D).Hence, our results indicated that *GRSF1*, *MAPT*, and *REC8* may contribute to tamoxifen resistance in breast cancer.

Discussion

Most breast cancers express ERα and belong to luminal A or luminal B subtypes [4]. Patients with these cancers exhibit a favorable prognosis owing to the efficacy of anti-estrogen drugs, such as tamoxifen. However, one-third of patients with luminal breast cancer eventually develop tamoxifen resistance, resulting in cancer progression and death. Therefore, identification of potential biomarkers and elucidation of the underlying molecular mechanisms of tamoxifen resistance in breast cancer are essential. In this study, using a method involving bioinformatics and experiment approaches, we finally screened out three hub genes involved in the development of tamoxifen resistance; these genes could be used as potential biomarkers to predict tamoxifen response and prognosis in ER-positive breast cancer.

In the current study, we first identified 1706 DEGs associated with tamoxifen resistance, including 859 upregulated genes and 847 downregulated genes. These DEGs were primarily enriched in functions such as protein targeting to the ER and pathways such as oxidative phosphorylation. As a hormonal transcription factor, ER regulates target genes to manipulate cell cycle progression and the endocrine response. The activity of ER is also regulated by multiple proteins, such as growth regulation by estrogen in breast cancer 1 (GREB1), transcription factors Ap-1, and FOXA1, thereby exerting differential biological functions in response to endocrine treatment[16-18]. On the other hand, many studies have shown that oxidative phosphorylation is closely correlated with carcinogenesis. Echeverria *et al.* reported that an oxidative phosphorylation inhibitor delayed residual tumor regrowth for neoadjuvant chemotherapy-resistant patients with breast cancer[19]. Moreover, Sansone *et al.*demonstrated that activation of oxidative phosphorylation promoted the development of hormone therapy-resistant disease [20].

WGCNA has emerged as a powerful method to obtain new insights into tumor biology. This approach can reveal co-expressed genes responsible for clinical features and can be used to screen out highly reliable and biologically significant modules and hub genes[11, 21, 22]. Our WGCNA clustered genes into 13 modules, of which the tan and blue modules were positively and negatively related to clinical traits, respectively. From this analysis, hub genes in these two modules were selected. Subsequent survival analysis showed that high expression of *GRSF1* predicted poor prognosis, whereas *MAPT* and *REC8* were associated with favorable survival outcomes in both all patients with breast cancer and patients with ER-positive breast cancer in TCGA dataset. Compared with previous studies [14, 15], we found modules and genes that were relevant to malignant phenotypes and favorable clinical features. More importantly, we validated the results using tamoxifen-sensitive and -resistant cell lines. Thus, *GRSF1*, *MAPT*,and *REC8* were the most promising candidate genes related to tamoxifen resistance.

GRSF1 was initially identified as an RNA-binding protein with high affinity for G-rich sequences. GRSF1 plays critical roles in maintaining mitochondrial functions, including mitochondrial translation, mitochondrial ribosome biosynthesis, and mitochondrial noncoding RNA binding[23, 24]. However, few studies have evaluated the roles of GRSF1 in cancer. Recent studies have reported that GRSF1 regulates microRNAs to facilitate oncogenic behaviors, including autophagy and metastasis, in cervical cancer

[25].Thus, combined with current results, these findings suggest that GRSF1 may function as a potential oncogene in cancer.

MAPT is a gene encoding Tau protein, which is implicated in the pathogenesis of a number of neurodegenerative disorders, including Parkinson's disease, Alzheimer's disease, and progressive supranuclear palsy[26]. Recent studies have suggested that high *MAPT* expression predicts better survival outcomes in pediatric neuroblastoma, low-grade glioma, breast cancer and clear renal cell cancer [27-29]. Additionally, Wang *et al.* reported that MAPT-hypermethylated tumors are associated with poor prognosis in patients with colorectal cancer[30]. Notably, MAPT plays an essential role in mediating paclitaxel or taxane resistance in various cancers. Several miRNAs have been found to determine paclitaxel chemosensitivity by regulating MAPT in gastric cancer and non-small cell lung cancer[31]. Rouzier*et al.* first identified MAPT as a predictor of the response to paclitaxel in breast cancer[32]. Moreover, clinical and in vitro studies have suggested that the expression level of MAPT is positively correlated with ER expression and is influenced by ER signaling. MAPT protein expression responds to tamoxifen in a dose-dependent manner[33, 34]. Consistent with these findings, our current results suggested that MAPT may participate in the response to tamoxifen in breast cancer.

REC8 is a key component of the meiotic cohesion complex and plays vital roles in the maintenance of chromosome stability and DNA damage repair[35]. Recent studies identified the novel role of *REC8* as a potential tumor-suppressor gene. *REC8*is hypermethylated in many cancers, including melanoma, thyroid cancer, and malignant gastrointestinal stromal tumor[36-38].Epigenetic inactivation of *REC8* predicts worse survival outcomes in patients with thyroid and gastric cancers. Moreover, Liu *et al.* reported that *REC8* is aberrantly methylated and robustly regulated by the phosphoinositide-3 kinase pathway in thyroid cancer[37]. The tumor-suppressive role of REC8 in gastric carcinogenesis is mediated by important genes implicated in the regulation of the cell cycle, proliferation, migration, and apoptosis[39]. Consistent with previous studies, our study also demonstrated that *REC8* may function as a favorable prognostic biomarker in breast cancer.

However, the current study still had some limitations, and additional studies are required in the future. First, tamoxifen resistance is controlled by a complicated regulatory network comprised of mRNAs, microRNAs, and long noncoding RNAs; however, only protein-coding genes/mRNAs were included in the analysis. Second, 155 tamoxifen-treated patients with breast cancer were included in the study. However, our model has not been verified in substantial clinical tissue samples and large prospective individual cohorts. Third, despite the validation of hub genes in breast cancer cell lines, we did not detect the protein level of each potential hub gene. Accordingly, further studies are needed to illustrate the mechanisms of gene networks involved in tamoxifen resistance.

Conclusion

Our current study identified gene networks and potential prognostic biomarkers using the systems biology-based WGCNA approach in patients with primary breast cancer treated with tamoxifen. The

genes in the tan and blue modules were positively or negatively associated with malignant behaviors, respectively. Additionally, based on survival analysis and preliminary in vitro experiments, we found three novel biomarkers related to tamoxifen sensitivity in ER-positive breast cancer; *GRSF1* was identified as a highly prognostic marker for cancer progression, whereas *MAPT* and *REC8* were found to predict favorable survival outcomes. However, further studies are still required to elucidate the exact molecular mechanisms and characterize the key genes functionally affecting tamoxifen sensitivity in patients with breast cancer.

Methods

Data collection and processing

The gene expression profile GSE9893 was obtained from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/) and evaluated using the GPL5049 platform [40]. The dataset GSE9893 compromised 155 tamoxifen-treated primary breast cancer samples, of which 52 cases developed recurrent disease (designated the tamoxifen-resistant group). Robust multiarray average background correction and $\log_2$ conversion were completed using the “affy” R package. Then, probes were mapped onto genes with Affymetrix annotation files. Genes matching with multiple probes were averaged as the expression level of the gene. Probes corresponding to multiple genes were deleted.

Analysis of DEGs

The R package “limma” R was employed to identify DEGs between relapse samples and relapse-free samples. The thresholds were set at values of $|\log_2 FC| > 1$ and $P < 0.01$. The results were visualized with the volcano plot R package.

GO and KEGG enrichment analyses

After obtaining DEGs related to tamoxifen sensitivity in breast cancer, the String database (https://string-db.org) was used to perform GO and KEGG pathway enrichment analyses to investigate the biological functions and pathways of tamoxifen resistance-related genes. GO analysis included three categories: biological processes (BP), cellular components (CC), and molecular functions (MF). The cut-off was set at an adjusted *P* value of less than 0.05.

Co-expression network construction by WGCNA

Co-expression networks were constructed using the WGCNA package in the R environment to explore the molecular mechanisms through which hub genes were related to tamoxifen resistance in breast cancer. First, we applied the sample network method to detect outlier samples. The soft threshold for WGCNA construction was selected such that the constructed network mainly included genes with strong correlations. Then, we transformed adjacency to the topological overlap matrix (TOM) to examine the connectivity of the network, followed by hierarchical clustering construction based on the TOM

dissimilarity to categorize genes with similar expression profiles into modules. Finally, analyses of module eigengene, gene significance, and module-trait relationships were performed to identify clinically significant modules.

Identification of hub genes

Hub genes were identified as highly interconnected genes in a module of WGCNA. In this study, tan and blue modules were considered key modules because they were closely related to the metastasis and recurrence of tamoxifen-resistant breast cancer. Hub genes were then screened out according to the absolute value of the Pearson's correlation. A maximum of 50 hub genes in modules were visualized using Cytoscape software.

Survival analysis

Survival analysis was performed with hub genes to further identify prognosis-associated genes using TCGA breast cancer dataset. All patients with breast cancer were classified into two groups according to the expression level of a particular gene (high versus low). Kaplan-Meier survival analysis was then employed to compare survival outcomes between groups using the Survival package in R. Patients with ER-positive breast cancer in TCGA database were utilized to validate the survival results of the candidate hub genes. Results with *P* values of less than 0.05 were considered statistically significant.

Cell culture

The human breast cancer cell line MCF-7 was purchased from the American Type Culture Collection (Manassas, VA, USA). Tamoxifen-resistant cells (MCF-7/TAM) were established by culturing MCF-7 cells in medium plus 1 μM4-hydroxytamoxifen (Sigma Aldrich, St. Louis, MO, USA) over 6 months. The cells were routinely maintained in Dulbecco's modified Eagle's medium with 10% fetal bovine serum at 37°C in an atmosphere containing 5% CO_2.

RT-qPCR

RNA extraction was performed using TRIzol reagent (TaKaRa Bio, Otsu, Japan), and cDNA was synthesized with a PrimeScript RT Reagent Kit (TaKaRa Bio) according to the manufacturer's protocol. Subsequently, qPCR was performed using SYBR Green (TaKaRa Bio) and a StepOnePlus Real-time PCR System (Applied Biosystems, Carlsbad, CA, USA). The primers used for RT-qPCR were: GRSF1 forward, 5'ACAGGGAAGAAATTGGTAATCG3' and reverse, 5' ACCATCGTCTACTGCCC TTTC 3'; MAPT forward, 5'AAAGACGGGACTGGAAGCG3' and reverse, 5'GAATC CTGGTGGCGTTGG 3'; and REC8 forward, 5'ATG AAAGCCTGGAATGGTTGG3' and reverse, 5'CCATATCCTGGAACGATTTGG3'. RT-qPCR was independently repeated three times.

Statistical analyses

SPSS 17.0 software (IBM, Armonk, NY, USA) was used to performed statistical analyses. Two-tailed Student's t-tests were used to detect differences between subgroups. Results with *P* values of less than 0.05 were considered significant.

Abbreviations

BP: biological process; CC: cell component; CCT8: chaperonin containing TCP1 subunit 8; COX7B: cytochrome c oxidase subunit 7B; ER: estrogen receptor;DEGs:differentially expressed genes;GO: Gene Ontology; GREB1: estrogen in breast cancer 1; GRSF1:G-rich RNA sequence binding factor 1;HER2: human epidermal growth factor receptor 2;JAK-STAT: Janus kinase/signal transducer and activator of transcription; KEGG: Kyoto Encyclopedia of Genes and Genomes;MAPT: microtubule associated protein tau; MF: molecular function;REC8: REC8 meiotic recombination protein; RT-qPCR:Reverse transcription quantitative polymerase chain reaction;TCGA: The Cancer Genome Atlas; TOM: Topological overlap matrix; WGCNA: Weighted gene co-expression network analysis

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All the authors have consented for the publication.

Availability of data and materials

The profiles of GSE9893 was obtained from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/).

Competing interests

No competing financial interests exist.

Funding

This study was supported by grants from the Natural Science Foundation of Zhejiang Province (LQ19H060002 and LQ19H160041) and Medical and Health Science and Technology Project of Zhejiang Province (2018KY089 and 2020KY143).

Authors' contributions

YYW and YJZ conceived and designed the experiments. YYW and XNG downloaded and analyzed the data. YYW performed in vitro experiments and wrote the manuscript. YJZ performed the critical revision

of the manuscript.

Acknowledgements

Not applicable.

Author details

[1]Department of Breast Surgery, The Second Affiliated Hospital, Zhejiang University School of Medicine, 88 Jiefang Road, 310009 Hangzhou, China.[2]Department of Orthopedic Surgery, The Second Affiliated Hospital, Zhejiang University School of Medicine, 88 Jiefang Road, 310009 Hangzhou, China.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7-34.
2. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet. 2018;391(10125):1023-75.
3. Zhang Y, Lv F, Yang Y, Qian X, Lang R, Fan Y, et al. Clinicopathological Features and Prognosis of Metaplastic Breast Carcinoma: Experience of a Major Chinese Cancer Center. PLoS One. 2015;10(6):e0131409.
4. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.
5. Early Breast Cancer Trialists' Collaborative G. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;365(9472):1687-717.
6. Early Breast Cancer Trialists' Collaborative G. Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. N Engl J Med. 1988;319(26):1681-92.
7. Ali S, Coombes RC. Endocrine-responsive breast cancer and strategies for combating resistance. Nat Rev Cancer. 2002;2(2):101-12.
8. Fan W, Chang J, Fu P. Endocrine therapy resistance in breast cancer: current status, possible mechanisms and overcoming strategies. Future Med Chem. 2015;7(12):1511-9.
9. Bianco S, Gevry N. Endocrine resistance in breast cancer: from cellular signaling pathways to epigenetic mechanisms. Transcription. 2012;3(4):165-70.
10. Szostakowska M, Trebinska-Stryjewska A, Grzybowska EA, Fabisiewicz A. Resistance to endocrine therapy in breast cancer: molecular mechanisms and future goals. Breast Cancer Res Treat.

2019;173(3):489-97.
11. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
12. Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM, et al. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. Proc Natl Acad Sci U S A. 2006;103(46):17402-7.
13. Di Y, Chen D, Yu W, Yan L. Bladder cancer stage-associated hub genes revealed by WGCNA co-expression network analysis. Hereditas. 2019;156:7.
14. Tang J, Lu M, Cui Q, Zhang D, Kong D, Liao X, et al. Overexpression of ASPM, CDC20, and TTK Confer a Poorer Prognosis in Breast Cancer Identified by Gene Co-expression Network Analysis. Front Oncol. 2019;9:310.
15. Jiang C, Wu S, Jiang L, Gao Z, Li X, Duan Y, et al. Network-based approach to identify biomarkers predicting response and prognosis for HER2-negative breast cancer treatment with taxane-anthracycline neoadjuvant chemotherapy. PeerJ. 2019;7:e7515.
16. Nilsson S, Makela S, Treuter E, Tujague M, Thomsen J, Andersson G, et al. Mechanisms of estrogen action. Physiol Rev. 2001;81(4):1535-65.
17. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat Genet. 2011;43(1):27-33.
18. Wu Y, Zhang Z, Cenciarini ME, Proietti CJ, Amasino M, Hong T, et al. Tamoxifen Resistance in Breast Cancer Is Regulated by the EZH2-ERalpha-GREB1 Transcriptional Axis. Cancer Res. 2018;78(3):671-84.
19. Echeverria GV, Ge Z, Seth S, Zhang X, Jeter-Jones S, Zhou X, et al. Resistance to neoadjuvant chemotherapy in triple-negative breast cancer mediated by a reversible drug-tolerant state. Sci Transl Med. 2019;11(488).
20. Sansone P, Savini C, Kurelac I, Chang Q, Amato LB, Strillacci A, et al. Packaging and transfer of mitochondrial DNA via exosomes regulate escape from dormancy in hormonal therapy-resistant breast cancer. Proc Natl Acad Sci U S A. 2017;114(43):E9066-E75.
21. Li B, Pu K, Wu X. Identifying novel biomarkers in hepatocellular carcinoma by weighted gene co-expression network analysis. J Cell Biochem. 2019.
22. Cai Y, Mei J, Xiao Z, Xu B, Jiang X, Zhang Y, et al. Identification of five hub genes as monitoring biomarkers for breast cancer metastasis in silico. Hereditas. 2019;156:20.
23. Antonicka H, Sasarman F, Nishimura T, Paupe V, Shoubridge EA. The mitochondrial RNA-binding protein GRSF1 localizes to RNA granules and is required for posttranscriptional mitochondrial gene expression. Cell Metab. 2013;17(3):386-98.
24. Jourdain AA, Koppen M, Wydro M, Rodley CD, Lightowlers RN, Chrzanowska-Lightowlers ZM, et al. GRSF1 regulates RNA processing in mitochondrial RNA granules. Cell Metab. 2013;17(3):399-410.

25. Sun Q, Yang Z, Li P, Wang X, Sun L, Wang S, et al. A novel miRNA identified in GRSF1 complex drives the metastasis via the PIK3R3/AKT/NF-kappaB and TIMP3/MMP9 pathways in cervical cancer cells. Cell Death Dis. 2019;10(9):636.
26. Bloom GS. Amyloid-beta and tau: the trigger and bullet in Alzheimer disease pathogenesis. JAMA Neurol. 2014;71(4):505-8.
27. Bonneau C, Gurard-Levin ZA, Andre F, Pusztai L, Rouzier R. Predictive and Prognostic Value of the TauProtein in Breast Cancer. Anticancer Res. 2015;35(10):5179-84.
28. Zaman S, Chobrutskiy BI, Blanck G. MAPT (Tau) expression is a biomarker for an increased rate of survival in pediatric neuroblastoma. Cell Cycle. 2018;17(21-22):2474-83.
29. Han X, Sekino Y, Babasaki T, Goto K, Inoue S, Hayashi T, et al. Microtubule-associated protein tau (MAPT) is a promising independent prognostic marker and tumor suppressive protein in clear cell renal cell carcinoma. Urol Oncol. 2020;38(6):605 e9- e17.
30. Wang C, Liu Y, Guo W, Zhu X, Ahuja N, Fu T. MAPT promoter CpG island hypermethylation is associated with poor prognosis in patients with stage II colorectal cancer. Cancer Manag Res. 2019;11:7337-43.
31. Cai Y, Jia R, Xiong H, Ren Q, Zuo W, Lin T, et al. Integrative gene expression profiling reveals that dysregulated triple microRNAs confer paclitaxel resistance in non-small cell lung cancer via co-targeting MAPT. Cancer Manag Res. 2019;11:7391-404.
32. Rouzier R, Rajan R, Wagner P, Hess KR, Gold DL, Stec J, et al. Microtubule-associated protein tau: a marker of paclitaxel sensitivity in breast cancer. Proc Natl Acad Sci U S A. 2005;102(23):8315-20.
33. Pentheroudakis G, Kalogeras KT, Wirtz RM, Grimani I, Zografos G, Gogas H, et al. Gene expression of estrogen receptor, progesterone receptor and microtubule-associated protein Tau in high-risk early breast cancer: a quest for molecular predictors of treatment benefit in the context of a Hellenic Cooperative Oncology Group trial. Breast Cancer Res Treat. 2009;116(1):131-43.
34. Ikeda H, Taira N, Hara F, Fujita T, Yamamoto H, Soh J, et al. The estrogen receptor influences microtubule-associated protein tau (MAPT) expression and the selective estrogen receptor inhibitor fulvestrant downregulates MAPT and increases the sensitivity to taxane in breast cancer cells. Breast Cancer Res. 2010;12(3):R43.
35. Nasmyth K. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. Annu Rev Genet. 2001;35:673-745.
36. Okamoto Y, Sawaki A, Ito S, Nishida T, Takahashi T, Toyota M, et al. Aberrant DNA methylation associated with aggressiveness of gastrointestinal stromal tumour. Gut. 2012;61(3):392-401.
37. Liu D, Shen X, Zhu G, Xing M. REC8 is a novel tumor suppressor gene epigenetically robustly targeted by the PI3K pathway in thyroid cancer. Oncotarget. 2015;6(36):39211-24.
38. Liu GB, Yan H, Jiang YF, Chen R, Pettigrew JD, Zhao KN. The properties of CpG islands in the putative promoter regions of human immunoglobulin (Ig) genes. Gene. 2005;358:127-38.
39. Yu J, Liang Q, Wang J, Wang K, Gao J, Zhang J, et al. REC8 functions as a tumor suppressor and is epigenetically downregulated in gastric cancer, especially in EBV-positive subtype. Oncogene.

2017;36(2):182-93.
40. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, et al. NCBI GEO: mining tens of millions of expression profiles–database and tools update. Nucleic Acids Res. 2007;35(Database issue):D760-5.

Tables

Table 1

Go enrichment analysis of up and downregulated DEGs

Category	Term	Gene ratio	adjusted *P*
Upregulated genes			
BP	GO:0072599-localization to endoplasmic reticulum	47/469	3.77E-41
BP	GO:0045047-protein targeting to ER	46/469	7.04E-41
BP	GO:0006614-protein targeting to membrane	45/469	9.98E-41
BP	GO:0006402-mRNA catabolic process	69/469	8.72E-39
BP	GO:0006413-translational initiation	50/469	2.94E-34
CC	GO:0022626-cytosolic ribosome	44/466	1.58E-42
CC	GO:0044391-ribosomal subunit	49/466	1.47E-35
CC	GO:0044445-cytosolic part	49/466	1.06E-31
CC	GO:0005840-ribosome	53/466	1.06E-31
CC	GO:0005925-focal adhesion	56/466	2.73E-24
MF	GO:0003735-structural constituent of ribosome	36/461	1.34E-20
MF	GO:0045296-cadherin binding	44/461	8.55E-18
MF	GO:0050839-cell adhesion molecule binding	51/461	4.30E-16
MF	GO:0019843-rRNA binding	13/461	1.43E-08
Downregulated genes			
BP	GO:0023061-signal release	33/440	4.37E-05
BP	GO:1903532-positive regulation of secretion by cell	30/440	4.71E-05
BP	GO:0060627-regulation of vesicle-mediated transport	33/440	0.00013
BP	GO:0046887-positive regulation of hormone secretion	15/440	0.0004
BP	GO:0051952-regulation of amine transport	15/440	0.0004
CC	GO:0043025-neuronal cell body	26/426	0.00178
CC	GO:0045211-postsynaptic membrane	20/426	0.00178
CC	GO:0045121-membrane raft	20/426	0.00213
CC	GO:0098857-membrane microdomain	20/426	0.00213
CC	GO:0097060-synaptic membrane	23/426	0.00244

MF	GO:0000987-proximal promoter sequence-specific DNA binding	29/432	0.00334
MF	GO:0042165-neurotransmitter binding	7/432	0.00334
MF	GO:0070405-ammonium ion binding	8/432	0.0028
MF	GO:0022890-inorganic cation transmembrane transporter activity	27/432	0.0034

Table 2

KEGG pathway enrichment analysis of up and downregulated DEGs

Category	Term	Gene ratio	adjusted *P*
Upregulated genes			
KEGG	hsa03010-Ribosome	47/333	1.84E-26
KEGG	hsa05012-Parkinson disease	27/333	2.70E-09
KEGG	hsa05010-Alzheimer disease	28/333	2.37E-08
KEGG	hsa04932-Non-alcoholic fatty liver disease (NAFLD)	26/333	2.37E-08
KEGG	hsa00190-Oxidative phosphorylation	22/333	1.05E-06
Downregulated genes			
KEGG	hsa04080-Neuroactive ligand-receptor interaction	30/275	0.00045
KEGG	hsa04972-Pancreatic secretion	13/275	0.00578
KEGG	hsa04925-Aldosterone synthesis and secretion	12/275	0.00909
KEGG	hsa04630-JAK-STAT signaling pathway	16/275	0.00909
KEGG	hsa04020-Calcium signaling pathway	17/275	0.01808

Figures

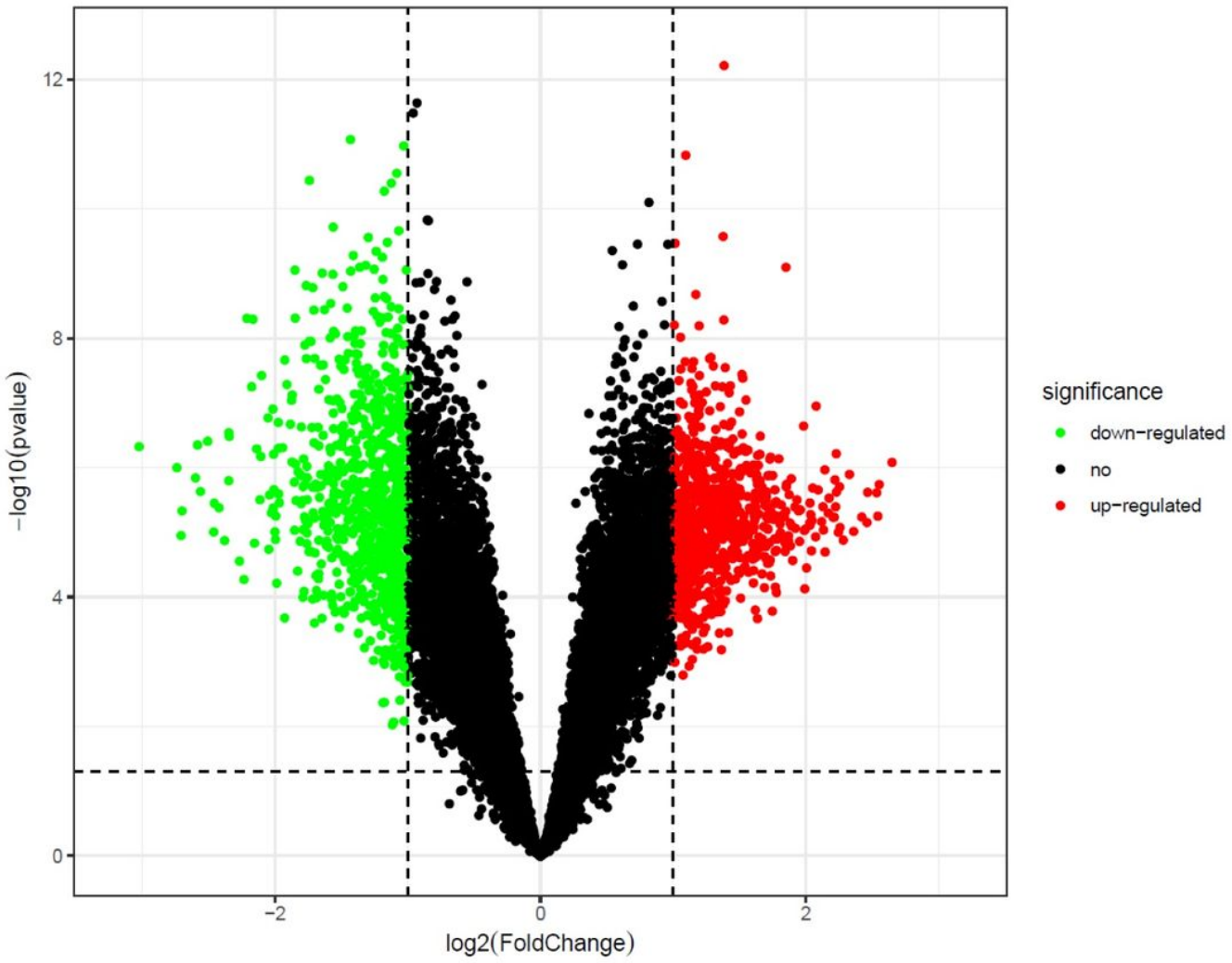

Figure 1

Volcano plot of the aberrantly expressedgenes between tamoxifen sensitive and resistant breast cancer groups. Green: high expression; Red: low expression; Black dots: the genes with expression of |log2FC| < 1 orP> 0.01.

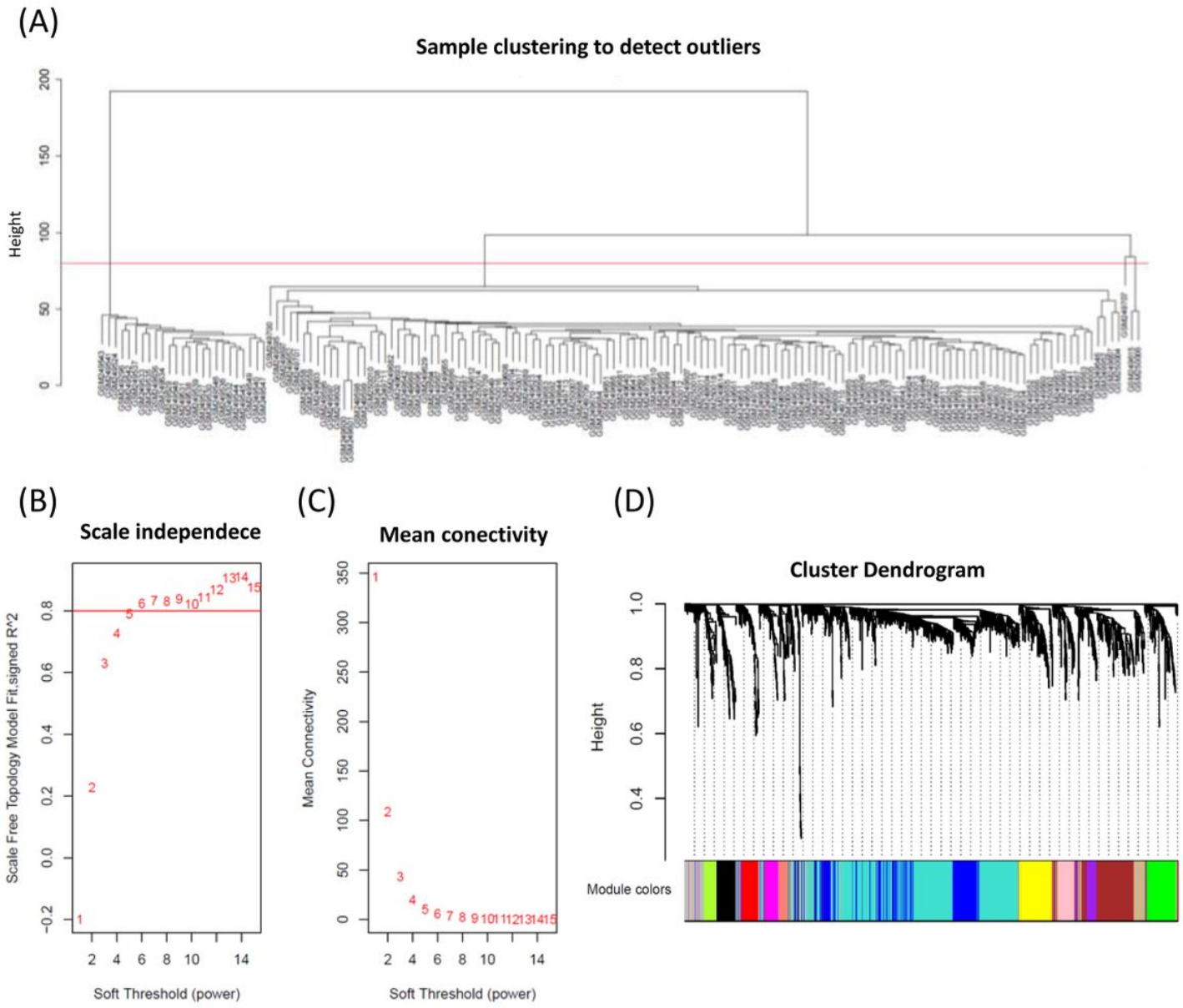

Figure 2

Clustering of samples and determination of soft-thresholding power in the WGCNA. (A)Samples clustering of GSE9893 to detect outliers. 28 samples were excluded. (B) Analysis of the scale-free fit index for soft-thresholding powers (β)from 1 to 15. (C) Analysis of the mean connectivity for variousβ values. β =6 was chosen for subsequent analyses. (D)A tree map of GSE9893 gene cluster. A total of 13 co-expression modules were constructed and displayed in different colors.

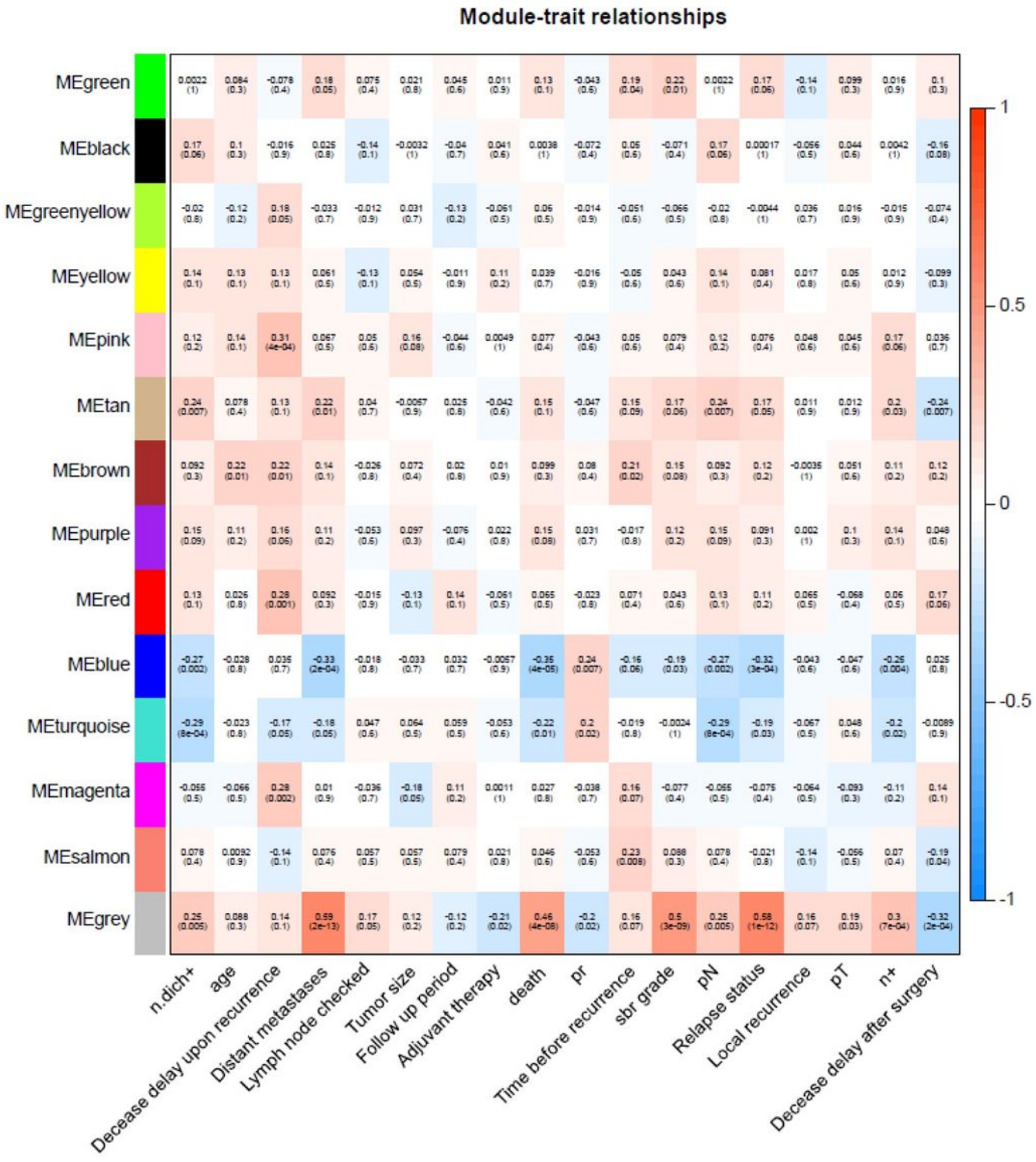

Figure 3

Identification of modules associated with clinical information. Each unit contains the corresponding correlation coefficient and Pvalue. The table is color-coded by correlation according to the color legend.

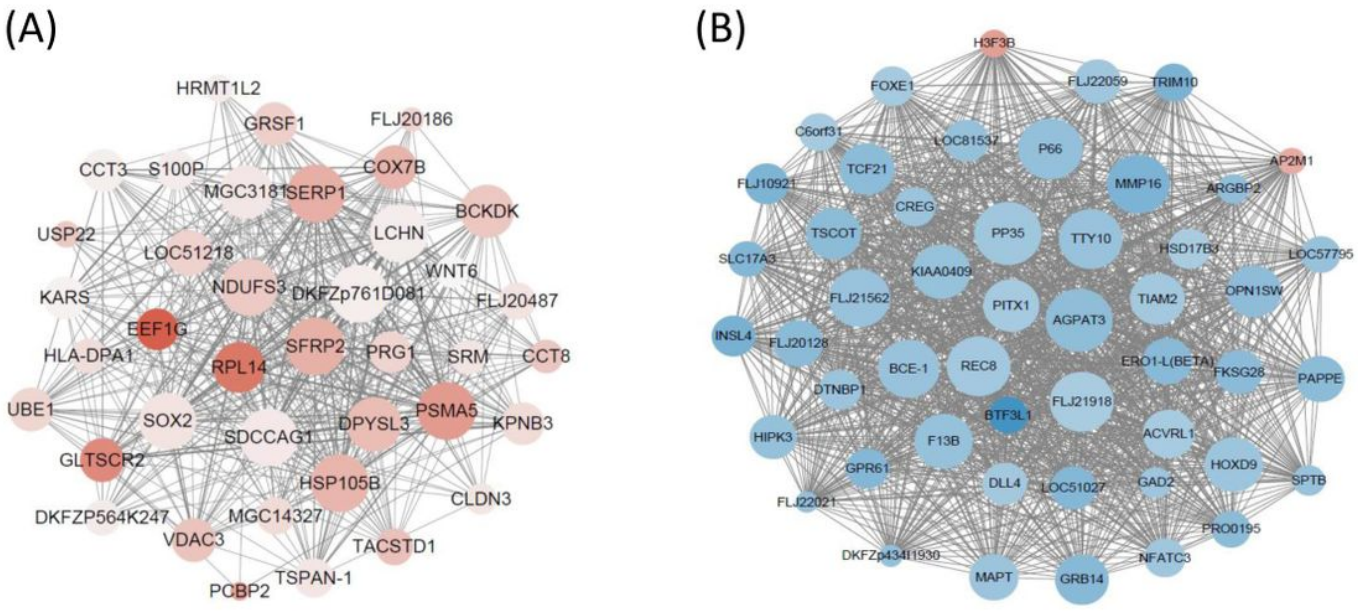

Figure 4

The visualization of hub genes in the tan (A) and blue modules (B). The red nodes represent up regulated genes and the blue nodes represent down regulated genes. Node size is correlated with the degree of connectivity for the corresponding gene.

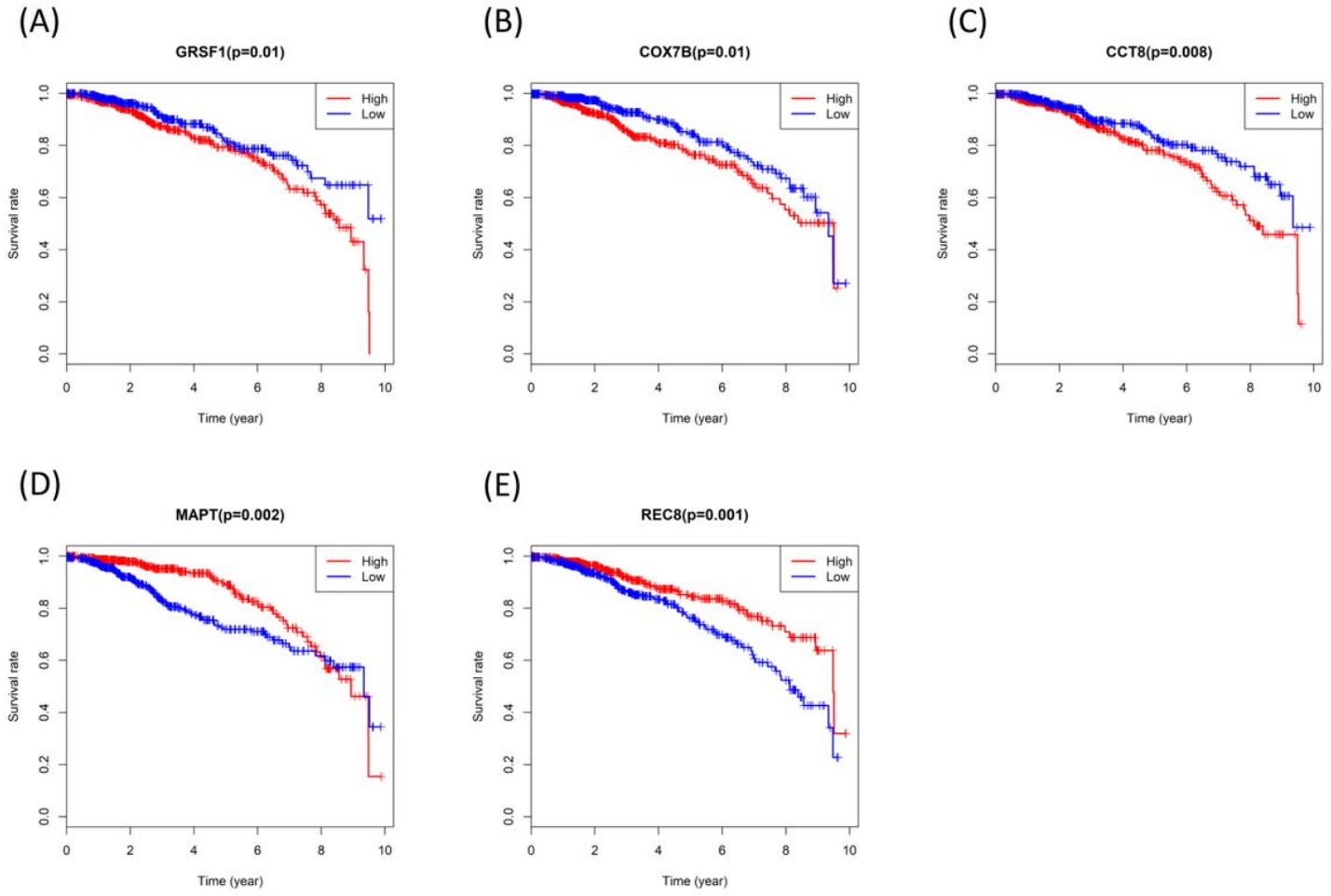

Figure 5

Survival analysis of prognosis-related hub genes in breast cancer patients from TCGA dataset. (A)GRSF1. (B)COX7B. (C)CCT8. (D)MAPT. (E)REC8.

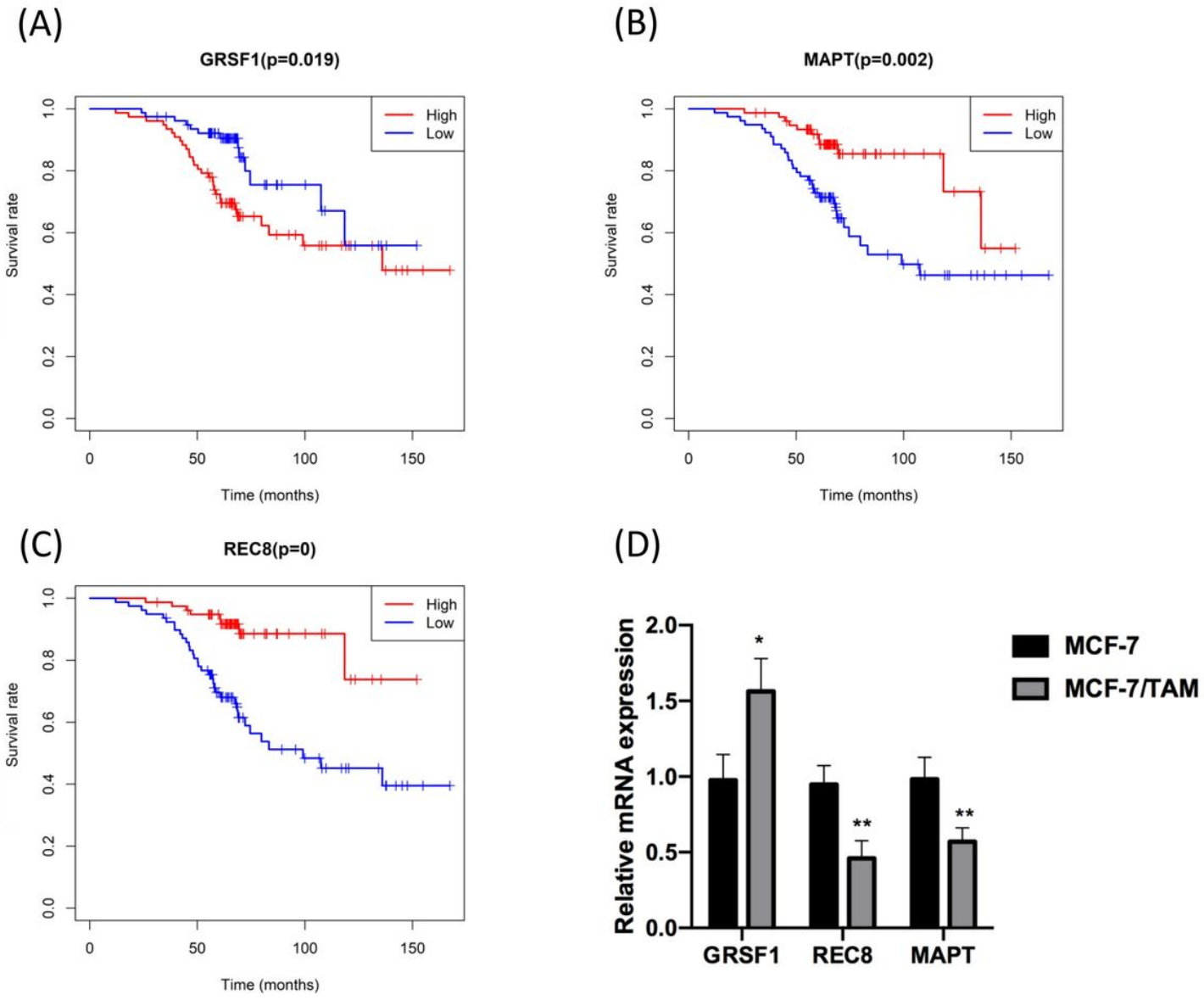

Figure 6

Validation of prognosis-related hub genes. (A-C) Survival analysis of hub genes in patients with ER-positive breast cancer from TCGA dataset. (D) RT-qPCR results of GRSF1, MAPT and REC8 in tamoxifen resistant/sensitive MCF-7 subclones. *P < 0.05 and **, P< 0.01 by two-tailed t test.